Lennart Bondesson

CHARACTERIZATIONS OF LOCATION AND SCALE PARAMETER FAMILIES OF DISTRIBUTIONS THROUGH OPTIMALITY PROPERTIES OF CERTAIN ESTIMATORS

Lund 1974

Lennart Bondesson

CHARACTERIZATIONS OF LOCATION AND SCALE PARAMETER FAMILIES OF DISTRIBUTIONS THROUGH OPTIMALITY PROPERTIES OF CERTAIN ESTIMATORS

Lund 1974

Copyright Lennart Bondesson 1974

Printed in Sweden

AV-centralen

Lund 1974

ACKNOWLEDGMENTS. I wish to express my deepest gratitude to

professor Gunnar Blom, for constant encouragement and support, and for help and advice concerning the presentation of the results;

fil mag Jacques de Maré and civ ing Holger Rootzén, who read earlier versions of four of the papers in detail and gave constructive criticism leading to many improvements;

professors Lars Gårding and Jaak Peetre, who gave valuable help on one of the papers;

fil lic Jan Lanke, fil dr Georg Lindgren, fil mag Bengt Ellenberger, and many other readers of my manuscripts, who pointed out both language errors and other errors. I am myself responsible for all remaining mistakes.

All the papers were partly written during my stay at the Mittag-Leffler Institute 1972-1973 which was made possible by financial support from the Mittag-Leffler foundation. I appreciate the stimulating discussions I had there with professor Lennart Carleson and many others.

Papers summarized in this dissertation

[A] Bondesson, L.: Characterizations of the gamma distribution. Teor. Veroyatnost. i Primenen. 18, 382-384 (1973)

[B] - - - - - - : Characterizations of the normal and the gamma distributions. Z. Wahrscheinlichkeitstheorie verw. Gebiete 26, 335-344 (1973)

[C] - - - - - - : When are the mean and the Studentized differences independent? To appear in The Annals of Statistics.

[D] - - - - - - : A characterization of the normal law. To appear in Sankhyā 1974.

[E] - - - - - - : Uniformly minimum variance estimation in location parameter families. To appear in The Annals of Statistics.

1. INTRODUCTION AND PRELIMINARY REMARKS

This dissertation is devoted to certain problems concerning uniformly minimum variance unbiased estimators (UMV-estimators) and best invariant estimators (Pitman estimators) of location and scale parameters. Generally speaking, we want to characterize the class of distributions for which a UMV-estimator or the Pitman estimator of such a parameter has some special form. We are also interested in finding all distributions for which there exists a non-constant parametric function that admits a UMV-estimator.

As a background for the discussions to follow, we consider an experimenter measuring some unknown physical quantity μ. The values he obtains, $x_1,\ldots,x_n$, are regarded as observations of a random variable (r.v.) X. Provided that there is no systematic error, $\mu = E[X]$. In order to estimate μ, the experimenter will generally, perhaps for lack of a better alternative, use the estimator $\mu^* = \bar{x}$, i.e. the arithmetic mean of the observations. There seem to be two ways of theoretically justifying the use of this estimator.

(i) If no a priori information on the distribution of X is available, $\bar{x}$ is the only unbiased estimator μ^* of μ which is symmetric in $x_1,\ldots,x_n$. Trivially, $\bar{x}$ is then also a UMV-estimator of μ.

(ii) If X is normally distributed, $N(\mu,\sigma)$, it is well known that $\bar{x}$ is a UMV-estimator of μ.

It is natural to ask whether the latter result remains valid for other so-called location-scale parameter families. This question started the investigations reported in this dissertation. More generally, three types of families of distribution functions are of interest.

(a) Location parameter family $F[x-\theta]$, $\theta \in R$.

(b) Scale parameter family $F[x/\sigma]$, $\sigma > 0$, $F(0) = 0$.

(c) Location-scale parameter family $F[(x-\theta)/\sigma]$, $\theta \in R$, $\sigma > 0$.

In Cases (a) and (c) the sample mean $\bar{x}$ is a UMV-estimator of its expected value if F is a normal distribution function (d.f.).

In Case (b) $\bar{x}$ has the same property if F is a gamma d.f. However, in each case we would like to know all d.f.'s F for which $\bar{x}$ is a UMV-estimator. Not surprisingly, the last type of family is much more difficult to study than the first two types.

We now turn to Pitman estimation. Let us first point out that best, with respect to squared-error loss, invariant estimators of θ and σ exist for each of the three types of families, as was shown by Pitman [B:7]*. We give the exact expressions of these estimators.

(a) Location parameter family $F[x-\theta]$

$$\theta_P^* = \bar{x} - E_0[\bar{x}|Y] ,$$

where $Y = (x_1-\bar{x},\ldots,x_n-\bar{x})$ and where the index 0 refers to $\theta = 0$.

(b) Scale parameter family $F[x/\sigma]$

$$\sigma_P^* = \bar{x}\cdot E_1[\bar{x}|Z]/E_1[\bar{x}^2|Z] ,$$

where $Z = (x_1/\bar{x},\ldots,x_n/\bar{x})$.

(c) Location-scale parameter family $F[(x-\theta)/\sigma]$

$$\theta_P^* = \bar{x} - s\cdot E_{(0,1)}[\bar{x}\cdot s|U]/E_{(0,1)}[s^2|U] ,$$

where $s^2 = \Sigma(x_i-\bar{x})^2/(n-1)$ and $U = \left((x_1-\bar{x})/s,\ldots,(x_n-\bar{x})/s\right)$.

In Case (a) θ_P^* is an unbiased estimator of θ. In Cases (b) and (c) both σ_P^* and θ_P^* are usually biased estimators of σ and θ, respectively.

Now one may ask for what distributions θ_P^* or constant$\cdot\sigma_P^*$ is equal to $\bar{x}$. These problems were formulated in the mid-1960's by Indian and Russian statisticians. In Case (a) the answer was given by the celebrated (but simple) Kagan-Linnik-Rao theorem: Provided that $n \geq 3$, $\theta_P^* = \bar{x}$ only for the normal distribution. This result was also the starting point for investigations in various directions by several Indian and Russian research workers. As a result there recently appeared a monograph: Characterization Problems in Mathematical Statistics (Wiley, 1973) written by the three above-mentioned authors. For us, Chapter 7 is of great interest.

In Case (b) Kagan and Rukhin [A:1] showed that if, for two different sample sizes ≥ 3, $\sigma_P^* = \text{constant}\cdot\bar{x}$, then F is a gamma d.f.

*) By [B:7] we mean the reference [7] in the paper [B].

It should be emphasized that as in Case (a) θ_P^* is an unbiased estimator of θ and as $\bar{x}$ is an invariant estimator, it follows that if $\bar{x}$ is a UMV-estimator, then $\theta_P^* = \bar{x}$. (The converse is true for $n \geq 3$ but not for $n = 2$.) In Cases (b) and (c) the analogous results do not hold as the Pitman estimators are biased. However, in Cases (b) and (c) there exist best, with respect to squared-error loss, unbiased invariant estimators of σ and θ. Thus, it is also of interest to find out when these are equal to $\bar{x}$. If in e.g. Case (b) $\bar{x}$ is a UMV-estimator of σ, then the best unbiased invariant estimator of σ is equal to $\bar{x}$. The converse is not true.

The Pitman estimators can be defined also for other losses than the quadratic one. Several authors have (with various degrees of success) tried to extend the Kagan-Linnik-Rao theorem to different types of such losses. See pp. 247-252 in the book by Kagan, Linnik, and Rao and also Fieger [4] and Klebanov [7].

It is of great historical interest to know that, as early as in the 19th century, Gauss [5] tried to find in Case (a) all d.f.'s F for which the maximum likelihood estimator of θ was equal to $\bar{x}$. By a not quite stringent method of proof, he found that only normal distributions were possible. Remarkably enough, this was the first time the Gaussian law appeared in Gauss' works. It is probably not very far from the truth to say that Gauss discovered the normal law in this way. Many papers have later been devoted to this problem, the best results being given by Csaszar [3] and Teicher [D:5]. In Case (b) the gamma distribution has also been characterized in the analogous way (see Teicher [D:5] and Ferguson [E:6]).

A general formulation of all the problems mentioned is provided by the following: When is the "best" estimator of θ (σ) given by $\bar{x}$ (constant$\cdot\bar{x}$)? Here "best" has different meanings. For instance, for Gauss "best" was equal to maximum likelihood.

Since a given parametric function admits a UMV-estimator only rarely, the general formulation may appear to the reader as not perfectly covering the UMV-case. However, as $\bar{x}$ is an invariant estimator for the three types of families, $\bar{x}$ is a UMV-estimator if and only if it is a locally best, with respect to squared-error loss, unbiased estimator at some arbitrary parametric

point. According to Barankin [2] and Stein [E:23] locally best unbiased estimators exist; this adds to the understanding of the problems.

We only mention that more general problems than the above ones are to find all distributions for which the "best" estimator of θ (σ) has some given form, not necessarily equal to $\bar{x}$. For instance, to what extent does the functional form of a Pitman estimator determine the distribution?

We now turn to UMV-estimation of a function of the location parameter θ in Case (a). The basic problem is to find all families for which there exists a non-constant parametric function that admits a UMV-estimator.

There are only four known different location parameter families admitting a complete sufficient statistic (one-dimensional, in fact), and for these families the celebrated Rao-Blackwell theorem shows that all estimable parametric functions (i.e. parametric functions that admit unbiased estimators) can be UMV-estimated. There are also general results due to Rao [B:8] and Bahadur [1] which assert that if all estimable parametric functions can be UMV-estimated, then there must exist a complete sufficient statistic. However, families $F[x-\theta]$ exist having no complete sufficient statistics, but admitting UMV-estimators of certain non-constant parametric functions. Here is an example.

Let $x_1,\dots,x_n$ be observations of the r.v. $X = X_1 + X_2$, where $X_1 \in N(\theta,1)$, where X_2 is a discrete r.v. taking the values $0, \pm1, \pm2, \dots$, and where X_1 and X_2 are independent. Let also g be a periodic function with period $1/n$. It is not hard to show that $g(\bar{x})$ is a UMV-estimator of its own mean value (see [E]).

Observe that this example is somewhat artificial since $E_\theta[g(\bar{x})]$ depends on n.

The basic problem seems to be extremely difficult to solve. In fact, there is little hope that a final solution can ever be found. We shall instead study simpler problems which essentially are of two different kinds.

(i) Given a statistic (or class of statistics), find all d.f.'s F

for which the statistic (or a member of the class) is a UMV-estimator.

(ii) Given an explicit d.f. F (or class of d.f.'s), find all statistics which are UMV-estimators of their expected values.

Problems of Type (i) have been discussed earlier. A simple problem of Type (ii) is to find all statistics which are UMV-estimators when F is normal. We leave to the reader to solve this problem. (Hint: use the Rao-Blackwell theorem and the uniqueness theorem for UMV-estimators.)

2. SUMMARY AND SUPPLEMENTARY REMARKS

We shall now summarize the contents of the papers [A],[B],[C], [D], and [E]. At the same time we also take the opportunity of pointing out some minor improvements of the results in [B]. An attempt is made to indicate briefly what methods have been used.

[A]: Scale parameter families $F[x/\sigma]$, $\sigma > 0$, $F(0) = 0$, are studied. It is shown that if, for some $n \geq 2$, the sample mean $\bar{x}$ is a UMV-estimator of its mean value, the distribution is gamma or degenerate. Provided that $n \geq 3$, the same conclusion holds if the Pitman estimator of σ or the best unbiased invariant estimator of σ is given by constant$\cdot\bar{x}$. It can be shown that the given lower bounds for n are the best possible ones.

The idea of the proofs is to compare the estimator $\bar{x}$ with estimators of the form

$$\bar{x} + \lambda x_1^{\zeta_1} x_2^{\zeta_2} \cdots x_n^{\zeta_n},$$

where λ is a real number and where the real numbers $\zeta_1,\ldots,\zeta_n$ have a fixed sum. The main tool is Cauchy's functional equation.

[B]: This paper is devoted to location-scale parameter families $F[(x-\theta)/\sigma]$, the d.f.'s F being continuous.

Besides the Pitman estimator θ_P^* of θ there exists a best unbiased invariant estimator θ_{UP}^* of θ. We present here the explicit expression for θ_{UP}^* since it is not given in the paper. We have

$$\theta_{UP}^* = \theta_P^* + B_n s \cdot E_{(0,1)}[s|U]/E_{(0,1)}[s^2|U],$$

where the constant B_n is chosen such that θ_{UP}^* is unbiased. To prove this one merely has to check that $E_{(0,1)}[\theta_{UP}^* sH(U)] = 0$ for all functions H of U such that $E_{(0,1)}[sH(U)] = 0$ and that is easily done.

As is well known, $\bar{x}$ is a UMV-estimator of θ if F is a normal d.f. with mean zero; hence it is also the best unbiased invariant estimator of θ. It is shown that this is true also if F is the d.f. of a r.v. $Y - pa$, where $Y \in \Gamma(p,a)$, $p > 0$, $a \neq 0$ (for the meaning of this, see p. 336 in the paper). (Notice that for $p \neq 1$ no complete sufficient statistic exists.) It is proved

that the converse is true: If, for some $n \geq 6$, $\bar{x}$ is the best unbiased invariant estimator of θ, then F is either a normal d.f. with mean zero or the d.f. of a variable $Y - pa$, where $Y \in \Gamma(p,a)$. (For the proof of this result the reader should also consult [C].) Furthermore, the Pitman estimator of θ is $\bar{x}$ for some $n \geq 6$ if and only if F is either a normal d.f. with mean zero or the d.f. of a variable $Y - pa - a/n$, where $Y \in \Gamma(p,a)$. It is not known whether $n \geq 6$ can be replaced by a weaker condition, e.g. $n \geq 5$. However, under the assumption that $\bar{x}$ is a UMV-estimator of θ it suffices to assume $n \geq 5$ in order to obtain the required conclusion.

It is also possible to characterize the normal distribution by the property that $\text{constant} \cdot s^2$ is a good estimator of σ^2. In fact, if the Pitman estimator of σ^2 or the best unbiased invariant estimator of σ^2 is given by $\text{constant} \cdot s^2$, then the distribution turns out to be normal. These results are proved under the assumptions that all moments are finite and that $n \geq 12$. However, utilizing a more refined technique, one can show that the moment condition is superfluous and that it suffices to let $n \geq 8$. Further, if we assume that s^2 is a UMV-estimator, it is sufficient to let $n \geq 6$.

Some similar results will be given in a forthcoming paper by Kagan and Zinger; only an abstract (Kagan and Zinger [6]) is available at present.

[C]: Let $U = \left((x_1-\bar{x})/s,\ldots,(x_n-\bar{x})/s\right)$ When X is normal or $X +$ constant is gamma distributed, it is well known that $E[\bar{x}|U] = \text{constant}$, $\bar{x}$ and U being independent. In this paper we show that, for $n \geq 6$, the constant regression property also characterizes these two distributions.

Some functions of U play important roles in the proof, namely

$$|x_1-x_2|^{it}|x_3-x_4|^{is}|x_5-x_6|^{-i(t+s)}$$

and

$$\left(|a_1x_1+a_2x_2+a_3x_3|+\tfrac{1}{2}(a_1x_1+a_2x_2+a_3x_3)\right)^{it}|a_1x_1+a_2x_2+a_3x_3|^{is}|x_4-x_5|^{-i(t+s)},$$

where $i = \sqrt{-1}$ and where the constants a_1,a_2,a_3 have the sum zero. An important tool is Cauchy's functional equation.

As pointed out above, the method utilized yields some improvements on the results in [B].

[D]: Location parameter families $F[(x-\theta)/\sigma_o]$, where both F and σ_o are fixed, are studied. The main result asserts that if F is sufficiently smooth and if the functional form of the Pitman estimator of θ does not depend on the actual value of σ_o, then F is a normal d.f. with mean zero. This result can be regarded as a partial generalization of the Kagan-Linnik-Rao theorem. A consequence is that for a location parameter family $F[x-\theta]$, F non-normal but smooth, the Pitman estimator of θ cannot be a linear combination of the order statistics. We also give an analogous characterization of the normal distribution by the property that the maximum likelihood estimator of θ is independent of σ_o. This result partly extends the above-mentioned result due to Gauss.

[E]: This paper is devoted to UMV-estimation in pure location parameter families $F[x-\theta]$. Fourier methods are intensively used.

The first part of the paper concerns UMV-estimation when only a single observation x is available. For an integer $N \geq 0$, let B_N denote the set of functions g for which $\sup|g(x)|/(1+|x|)^N < \infty$. The meaning of the concept B_NUMV is then also obvious. Let φ be the characteristic function of F, i.e. $\varphi(\zeta) = \int \exp\{i\zeta x\} dF(x)$. Further, for a function g, we let $S(g)$ denote the support of its Fourier transform. The following result is established.

Provided that F is absolutely continuous with at least $2N + 1/2$ finite moments and provided that g belongs effectively to B_M $(M \leq N)$ (for the meaning of this, see the paper), $g(x)$ is a B_NUMV-estimator if and only if both

$$S(g) \subset \{ \zeta \in R ; \varphi(\eta-\zeta) = 0 \text{ for all } \eta \in R \text{ such that } \varphi(\eta) = 0 \}$$

and every real zero of φ has multiplicity at least $N+M+1$.

Another to some extent surprising result is that if the tail of F tends to zero rapidly enough, then $g(x)$, where g is aperiodic and not increasing too fast, is a UMV-estimator only for the normal distribution.

According to a well-known UMV-criterion, a statistic is a UMV-estimator of its mean value if and only if it is uncorrelated with every unbiased estimator of zero. Besides this criterion the proofs of the results above utilize the fact that most unbiased estimators

of zero can be approximated by finite linear combinations of statistics of the form $\exp\{i\zeta x\}$, where $\varphi(\zeta) = 0$.

The second part of this paper treats sample sizes ≥ 2. A basic necessary (and nearly sufficient) condition for an estimator to be UMV is established. The proof of the condition utilizes the fact that the statistics

$$\exp\{i\zeta_1 x_1 + i\zeta_2 x_2\}/\varphi(\zeta_1)\varphi(\zeta_2) - \exp\{i\eta_1 x_1 + i\eta_2 x_2\}/\varphi(\eta_1)\varphi(\eta_2),$$

where $\zeta_1 + \zeta_2 = \eta_1 + \eta_2$, and

$$\exp\{i\zeta_1 x_1 + i\zeta_2 x_2\},$$

where $\varphi(\zeta_1) = 0$ or $\varphi(\zeta_2) = 0$, both are unbiased estimators of zero. It also seems as though, for $n = 2$, most unbiased estimators of zero in some way can be approximated by finite linear combinations of statistics having the above forms.

The condition is the Fourier form of an old operator condition given by Barankin [2] and Stein [E:23]. An application of the condition shows that for the uniform distribution no non-constant parametric function can be UMV-estimated. A generalization of this result is also given.

Three UMV-characterizations are derived. The normal distribution is shown to be the only one admitting, for some $n \geq 2$, a UMV-estimator of the form $g(\bar{x})$, where g is aperiodic and not increasing too fast. The aperiodicity assumption is necessary as the example in the introduction shows. One should here recall the Kagan-Linnik-Rao theorem which asserts that, for $n \geq 3$, $\theta_P^* = \bar{x}$ also only for the normal law. If a statistic $g(x_1) + \ldots + g(x_n)$ is a UMV-estimator, then, under slight assumptions, F is normal, the d.f. of the logarithm of some power of a gamma distributed r.v. (the corresponding family admits a complete sufficient statistic), or a lattice d.f. It is shown that no polynomial can be a UMV-estimator, unless the distribution is normal or has finite support.

As mentioned before, a UMV-estimator T is not necessarily also a sufficient statistic. However, it is proved that this is the case if (i) T has moments of all orders which uniquely determine the distribution of T; and (ii) the mapping: $(x_1,\ldots,x_n) \to (T, x_1-\bar{x}, \ldots, x_n-\bar{x})$ is injective.

UMV-estimation of the location parameter itself is studied. If the tail of F tends to zero rapidly enough, e.g. if the measure dF has compact support, it turns out that no UMV-estimator of θ exists unless the distribution is degenerate.

REFERENCES

The references given below complement the references in the papers.

[1] Bahadur, R.R: On unbiased estimates of uniformly minimum variance. Sankhyā 18, 211-224 (1957)

[2] Barankin, E.W.: Locally best unbiased estimates. Ann. Math. Statist. 20, 477-501 (1949)

[3] Csaszar, A.: Sur une caractérisation de la répartion normale de probabilité. Acta Math. Hung. 7, 359-382 (1956)

[4] Fieger, W: Eine statistische Charakterisierung der Normalverteilung. Z. Wahrscheinlichkeitstheorie verw. Gebiete 19, 330-344 (1971)

[5] Gauss, C.F.: Theoria motus corporum coelestium, Werke, Liber II, Sectio III, 240-244

[6] Kagan, A.M., Zinger, A.A.: Some statistical problems for location and scale parameter families of distributions. International Conference on Probability Theory and Mathematical Statistics in Vilnius 1973. (Russian)

[7] Klebanov, L.B.: The admissibility of the sample mean as an estimate of the location parameter for non-quadratic risk-functions. Theory Probab. Appl. 18, 329-339 (1973)

AV-Centralen, Lund 1974